CHAPTER 5
THE FORMATION OF STARS

5.1 Introduction

As with so many astronomical phenomena, the life cycle of a star takes place over a timescale that appears infinitely long in comparison with a human lifetime, or even with the entire duration of recorded human history of perhaps several thousand years. Zoologists or botanists are usually able to study the complete life cycles of the animals or plants in which they are interested, but astronomers are not afforded that luxury where the stars are concerned. The changes that occur in stars are, with a few notable exceptions, much too slow to be observed. The evolutionary pattern has to be deduced by observing a wide range of stars at different stages of their lives, and combining these observations with theory and models based on the laws of physics as determined in other environments. The fact that a credible model of the entire life cycle of most types of star has been developed can surely be regarded as one of the triumphs of 20th century astrophysics. This doesn't mean, however, that all the problems are solved – far from it. There is still much to understand.

Before we consider the physical processes and energy sources that govern the structure of stars we will consider the process of star formation.

5.2 Conception

Where should we be looking to find starbirth taking place? The fundamental idea is that stars are born from more widely dispersed gas which condenses because of the gravitational attraction of the gas on itself. The British astronomer Sir James Jeans was responsible for placing this idea on a firm mathematical footing. He showed how an extended mass of gas could be unstable because of the gravitational forces within it and we will follow his arguments in Section 5.3. We have seen (Section 3.2.4 and Figure 0.1) that the distribution of stars within the Milky Way is far from even. We have also seen, in Section 4.3, that the apparent lack of stars in certain directions arises because our view of the stars is obscured in those directions by

JAMES JEANS (1877–1946)

James Jeans (Figure 5.1) showed his scientific and literary prowess at an early age, writing a handbook on clocks at the age of nine, including instructions on how to make a clock from pieces of tin! He graduated from Cambridge in 1900 and after academic posts at Princeton and Cambridge he devoted his time to private research and writing from 1912. He made contributions to the kinetic theory of gasses and quantum physics, and formulated the Rayleigh–Jeans law (which describes the spectral energy distribution of the long wavelength end of a black-body spectrum), before turning to astrophysics. His studies of fluids led him to consider the origin of the Universe, stars and the Solar System. His belief that matter was being continuously created in the Universe (a forerunner of the Steady State Theory) was not developed further as he concentrated on broadcasting and popular books after his knighthood in 1928.

Figure 5.1 James Jeans. (Royal Astronomical Society)

intervening clouds of gas (consisting mainly of hydrogen) and dust (about 1% of the mass of the gas and consisting mainly of heavy elements and their compounds). Detailed study of the spectra of stars reveals that more tenuous matter is spread throughout interstellar space. Before we look at the process of star formation we will examine further the properties of different components of the interstellar medium (ISM) and the regions most likely to be the birthplace of stars.

5.2.1 The interstellar medium

The temperature and the number density are two important physical parameters of a gas, and we find that these two quantities have a wide range of values in the ISM. The number density n is the number of particles per unit volume and so has SI units of m^{-3}. Atoms can be neutral, ionized, or combined in molecules or in dust. Because of the predominance of hydrogen, to a sufficient approximation we can sometimes take n to be the number of hydrogen atoms per cubic metre, n_H. However, even when we ignore the other elements, n is not always the same as the number of hydrogen atoms per cubic metre. For example, if all the hydrogen is present in the molecular form, H_2, then the number of separate particles (molecules) per cubic metre is about $n_H/2$.

- What is the particle number density if all the hydrogen is ionized?
- About $2n_H$ (the electron is a separate particle when the hydrogen is ionized so there are n_H hydrogen nuclei and n_H electrons).

Only if hydrogen is present mainly as un-ionized (neutral) atoms will the particle number density, n, be about the same as the hydrogen atom number density n_H. However, n_H always gives a good measure of the mass density ρ. Because hydrogen nuclei are the main contribution to the mass of any sample of ISM, $\rho \sim n_H m_H$, where m_H is the mass of the hydrogen atom, little changed by ionization, and unchanged by combination in molecules.

Figure 5.2 shows the typical number densities and temperatures found in the various regions of the ISM. First, note the enormous range of number densities, from a minuscule value of about $100\,m^{-3}$, to the much greater value of about $10^{17}\,m^{-3}$ (although this is still a very low density as you will see). Second, note that the temperature range is also enormous, from less than 10 K, to several million kelvin – comparable with temperatures found in stellar interiors. However, within these wide ranges, the various temperatures and number densities are not equally represented. Thus, the names given in Figure 5.2 are not of arbitrary subdivisions, but correspond to locations on the diagram where the measured temperature and number density values tend to concentrate. It is because there are such concentrations that temperature and number density provide a useful basis on which to define some of the different types of region. Table 5.1 lists the properties of the different regions.

The intercloud media (hot and warm) account for most of the volume of the ISM, and together form a low-density, optically transparent, widespread matrix in which the other types of region are embedded. Each of these other types is present as a large number of separate objects: typical sizes are given in Table 5.1.

The **hot intercloud medium** is very widespread. You might therefore be wondering why it doesn't blaze down on us from the night sky. The reason is its extremely low density, coupled with its highly ionized state.

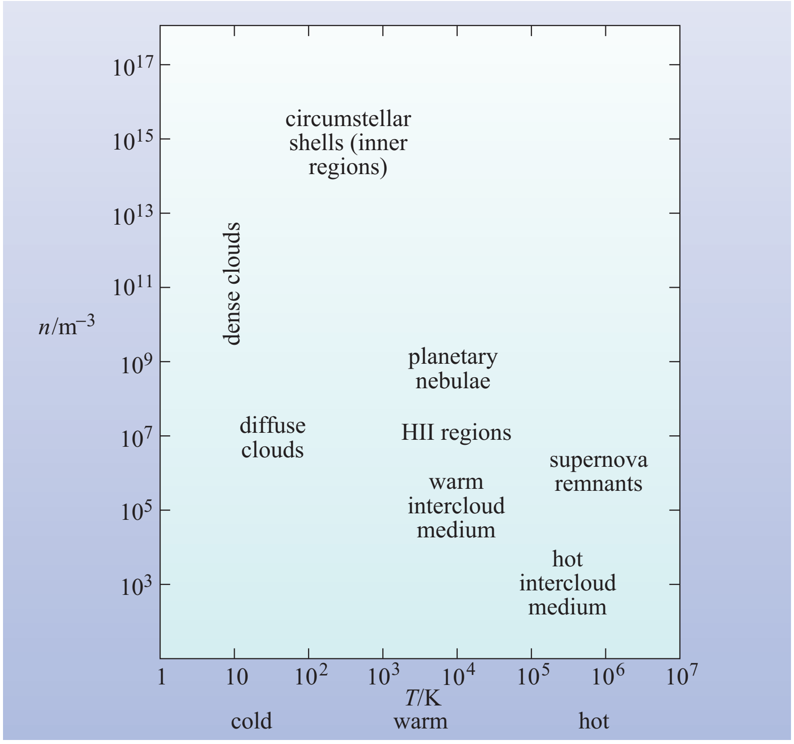

Figure 5.2 Various types of region in the ISM, distinguished on the basis of the number density n of particles, and the temperature T. The atoms can be neutral, or ionized, or combined in chemical compounds.

Table 5.1 Some features of the various types of region in the ISM.

Type of region	Fraction of the ISM (%)[a]		Typical size/pc[c]	Typical mass/$M_\odot$	Predominant form of hydrogen	Abundance of molecules
	By volume[b]	By mass				
hot intercloud medium	~60	≤0.1	–	–	H^+	very low
warm intercloud medium	~30	~20	–	–	H^+ or H	very low
diffuse clouds	~3	~30	~3 to ~100	1 to 100	H or H_2	diatomic molecules common
dense clouds	≤1	~45	~0.1 to ~20	1 to 10^4	H_2	molecules common, even large ones
HII regions	~10	~1	~1 to ~20	10 to 10^4	H^+	very low
circumstellar shells	negligible	negligible	≤1	≤1	H or H_2	diatomic and small molecules common
planetary nebulae	negligible	negligible	≤2	≤1	H^+	very low
supernova remnants	[d]	negligible	≤1000	~3	H^+ or H	very low

[a] These percentages are only rough estimates, so do not sum to 100%.

[b] The total volume of the ISM is taken to be a disc with a diameter roughly that of the spiral arms of the Milky Way, i.e. 30 000 pc, and a thickness of 300 pc.

[c] These are typical distances across a region (e.g. diffuse and dense clouds are usually irregularly shaped, and are often more sheet-like than spherical). For roughly spherical regions the size is roughly the diameter.

[d] The volume is included in the hot intercloud medium (see text).

- Would you expect any radiation from electronic transitions (see Section 1.3.2) in a low-density, fully ionized hydrogen gas?
- No. Low-density gases can produce emission lines when illuminated by a hot source but hydrogen has only one electron and there are therefore no bound electrons in ionized hydrogen to change energy state and produce a spectral line.

The hot intercloud medium consists mostly of fully ionized hydrogen, and the low density means that in any case there is little material to radiate. The low density also means that it is highly transparent, and so the medium does not obscure anything lying in or beyond it. For much the same reasons the **warm intercloud medium** is also not very apparent.

By contrast, the other types of region are much more readily detected, often at a variety of wavelengths. These other types of region can be subdivided into those that are associated with individual stars at or near the end of their lives (planetary nebulae, supernova remnants and circumstellar shells that you will hear more about in Chapter 8), and those that are not (diffuse and dense clouds and HII ('aitch 2') regions described below).

Diffuse clouds are cold regions of moderate density such that stars can be seen through them at visible wavelengths, and consequently they are not apparent to the unaided eye. Hydrogen is present in both atomic (H) and molecular (H_2) forms, and a number of other simple molecules are also found. The clouds are often mapped, and otherwise investigated by various techniques. Atomic hydrogen can be detected through an emission line at a wavelength of 21 cm. Molecular hydrogen cannot be mapped directly, but the CO molecule, which usually occurs with H_2, does emit spectral lines in the microwave band. Finally, the diffuse clouds contain dust that can be mapped from its far infrared emission.

Dense clouds are as cold as or colder than diffuse clouds, and, unsurprisingly, are denser! Unless it is very thin in the direction of our line of sight, a dense cloud is opaque at visible and UV wavelengths, because of extinction by the dust in it. It is then seen by its obscuration of the stars beyond, and is often called a dark cloud. Such obscuration is apparent in Figure 4.14, which shows the Coal Sack, a dark cloud near the Southern Cross, readily seen with the unaided eye from any latitude south of about 25° N. The gas in dense clouds consists largely of molecules. In addition to H_2 there are a great variety of less abundant but nevertheless important molecules. Moreover, extinction by dust is far less severe at longer wavelengths, and so mapping, and other investigations, often rely on the (collisionally excited) microwave emissions from molecules such as HCN, OH, CS and CO. Figure 5.3 shows a map of CO emission in and around Orion: the more copious the emission, the greater the amount of material. Roughly speaking, the regions apparent in this map are dense clouds, and, though unseen here, they are fringed in most places by diffuse clouds. Typically, the two types of region are closely associated.

An **HII region** is a low-density cavity created inside a dense cloud by one or more hot, bright stars (of spectral class O or B) which emit large amounts of ultraviolet radiation. This ionizes the hydrogen gas in their proximity. For historical reasons, the term HII is used to indicate hydrogen in its ionized form.

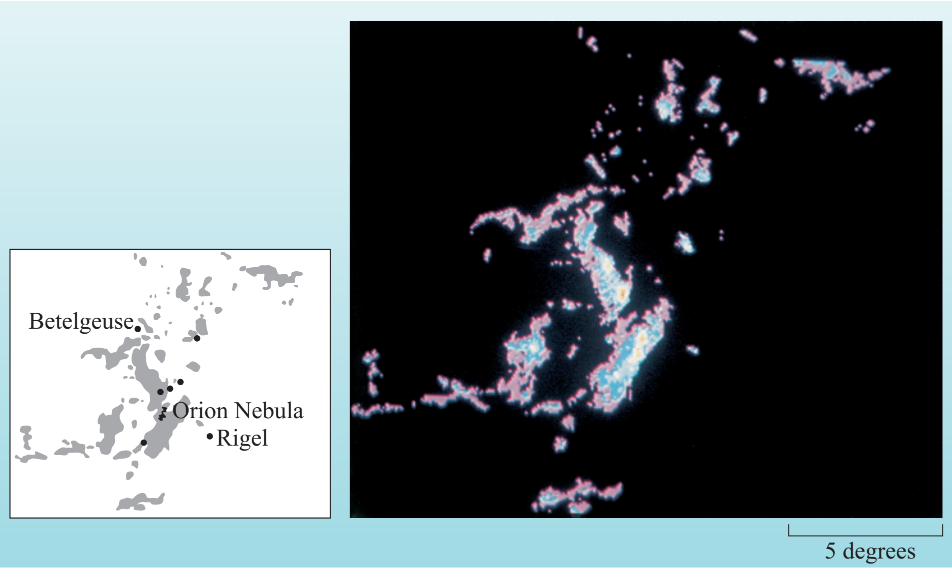

Figure 5.3 The constellation of Orion as seen at radio wavelengths emitted by CO molecules. The density of the gas is indicated by the colour, black where it is undetectable, violet where it is least dense, then through blue to white where it is most dense. Although covering a similar region of sky as Figure 3.1, the view is very different. Molecules pervade the cool regions of the interstellar medium whereas the visible image is dominated by stars or the effect of starlight on the surrounding gas and dust (see Section 4.3). (Image: R. Maddalena/NRAO)

(An alternative name for this type of region is 'diffuse nebula', but this is easily confused with the very different diffuse clouds, so we won't use it.) The hot ionized hydrogen has a greater pressure than the colder un-ionized hydrogen so it expands outwards into the colder material, creating a low-density region inside the cold dense cloud.

In many cases, the HII region has burst through the surface of the dense cloud, and is seen at visible wavelengths. This is the case for the Orion Nebula shown schematically in Figure 5.4. The nebula is lit up by the intense radiation from the young stars (four of which at the centre of the nebula form the Trapezium shown in Figure 6.14).

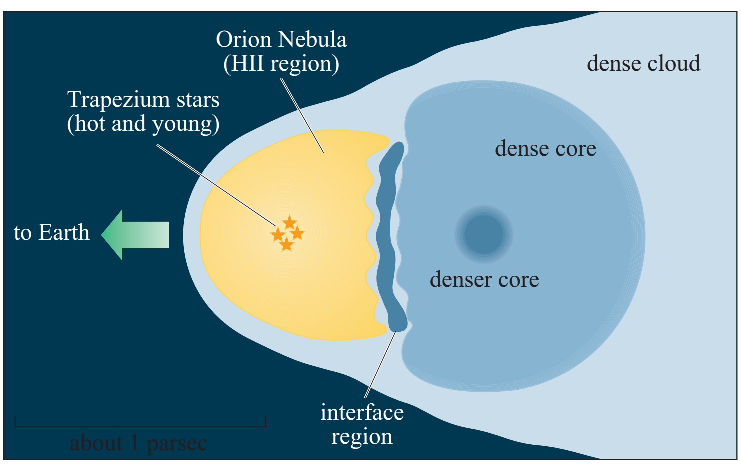

Figure 5.4 A cross-section through the Orion Nebula (simplified).

So, what produces the characteristic red colour of HII regions? The key point here is that not *all* the gas in an HII region is ionized all of the time. Occasionally a free electron and a hydrogen nucleus (proton) can reform a neutral hydrogen atom – a process called **recombination**. The electron is typically captured into a high-energy orbit and then cascades downward through the atom's energy levels emitting photons as it does so. The red coloration from HII regions is due to a transition from energy level $n = 3$ to $n = 2$ (see Figure 3.20). This transition gives rise to the Hα line at a wavelength of 656.3 nm.

Any region in which hydrogen is predominantly in molecular form is called a molecular cloud. Thus *all* dense clouds are molecular clouds, and so too are some diffuse clouds.

Many diffuse clouds have dense clouds moving around inside them (Figure 5.3). Moreover, within a dense cloud, whether inside a diffuse cloud or not, there are often even denser regions called cores and clumps, with masses in the range from $\sim 0.3M_\odot$ to $\sim 10^3 M_\odot$. The largest of such clouds are called **giant molecular cloud complexes**, or GMC complexes for short. They have dimensions up to ~100 pc and masses up to $\sim 10^6 M_\odot$. They are being increasingly recognized as the fundamental cloud structure in the ISM, rather than the individual diffuse or dense clouds. Figure 5.5 illustrates the hierarchical structure of a GMC complex, and Figure 5.3 shows that the Orion Nebula (Figure 4.13), which you have already met, is just one small part of a larger complex.

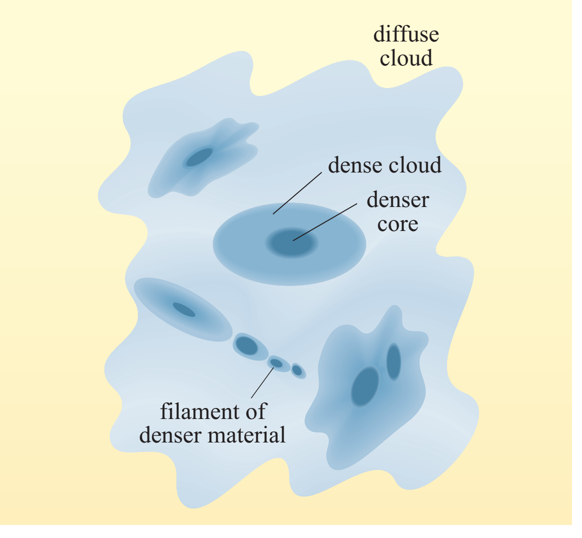

Figure 5.5 The hierarchical structure of a giant molecular cloud complex.

At this point we should put the term 'dense' into context.

QUESTION 5.1

At the surface of the Earth, the Earth's atmosphere has a temperature of about 300 K, and a density of about 1 kg m^{-3}. Where, in Figure 5.2, do these physical conditions lie? You can assume the atmosphere is entirely made up of its major constituent, nitrogen molecules, N_2.

So, comparing a cubic metre of material from a dense interstellar cloud with a cubic metre of the Earth's atmosphere at sea-level, you have found that the sample of atmosphere has about 10^{15} (a thousand million million) times more gas molecules and atoms than the sample of the same volume from the dense cloud. Even if you took a similar sized sample from what we would describe as 'a good laboratory vacuum', you would still find that a sample from the vacuum would have a factor of about 10^6 more particles than the sample from the dense cloud. However, the clouds are large, and if you multiply the large volumes by the low densities, you do end up with a significant amount of material. The dense clouds make up perhaps 45% of the total mass in the interstellar medium. It is in the dense clouds where it appears that conditions are particularly favourable for stars to form.

5.2.2 Dense clouds – the interstellar nursery

Let's look at the evidence that supports the belief that dense clouds are the site of starbirth.

Some young star clusters seem to be surrounded by the remnants of the original cloud from which they formed. Figures 3.1 and 4.13 show the Orion Nebula. It is visible to the naked eye or through binoculars as a haze in the constellation of Orion. Observations with a telescope show this region to be one of the most visually magnificent in the sky. There is plenty of evidence to suggest that star formation took place here very recently, that is it is recent on an astronomical timescale. For example, if you look at an H–R diagram of the stars in the central region of the nebula (Figure 5.6), you see that they mostly lie above the main sequence.

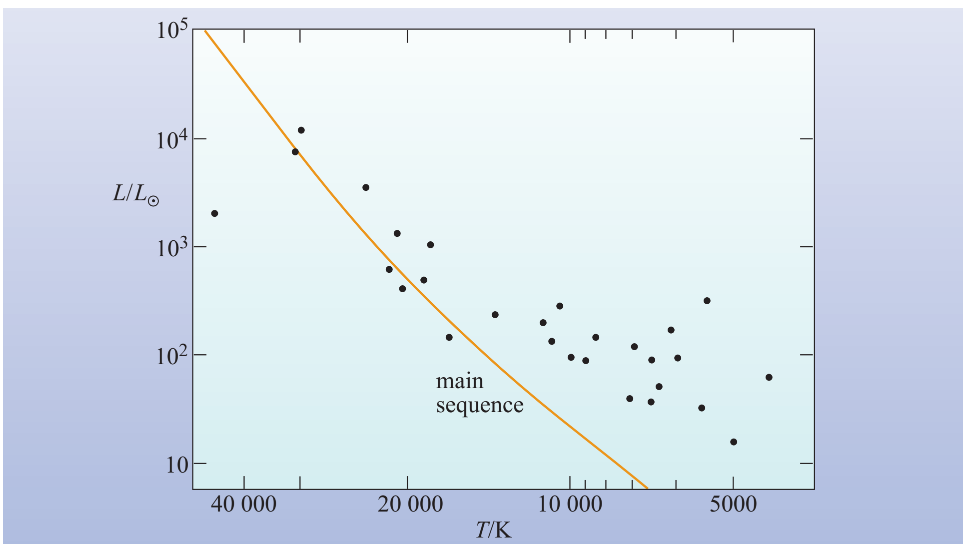

Figure 5.6 H–R diagram for stars in the Orion Nebula, showing that most of them fall above the main sequence. (Note that this H–R diagram covers only part of the range of L and T shown in Figure 4.5.)

■ What does this suggest about the stars in this region?

❑ You saw in Chapter 4 that stars spend most of their life, maybe 90% of it, on the main sequence. The fact that the stars associated with the Orion Nebula aren't on the main sequence suggests therefore that they are at either the beginning or the end of their lifetime.

Figure 5.7 The star cluster NGC 2264 is formed of young stars illuminating the remnants of the dense cloud from which it formed. (D. Malin/AAO)

All other evidence points to these stars being very young. What interests us here, though, are the vast clouds of glowing gas and obscuring dark clouds of dust and gas. These are almost certainly the remnants of the dense cloud from which the young stars in the nebula formed.

Another example is in Figure 5.7 where the star cluster NGC 2264 (object number 2264 in the 'New General Catalogue of Nebulae and Clusters of Stars' which was an updated version of John Herschell's (see Figure 3.11) 'General Catalogue of Nebulae and Clusters' and published in 1888!) is seen to be intimately associated with the remnants of a dense cloud. Most of these stars also lie above the main sequence on the H–R diagram, though they seem to be rather older than those in the Orion Nebula. Theoretical estimates yield ages of about 5×10^6 years for the stars in NGC 2264, and about 10^6 years for those in the Orion Nebula.

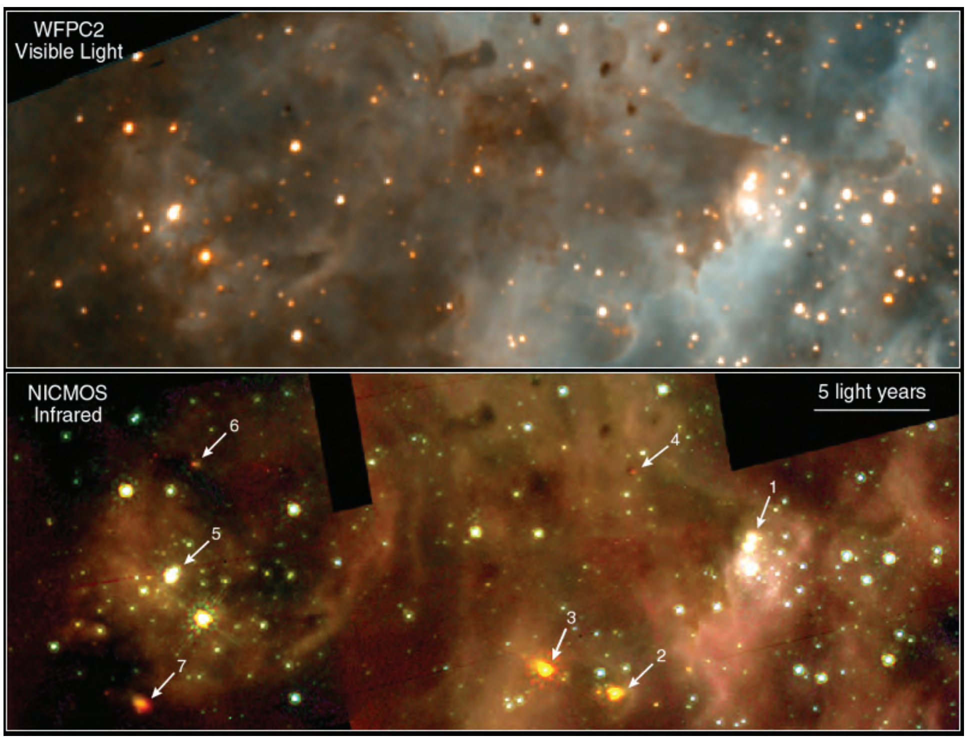

Figure 5.8 Visible light (top) and infrared (bottom) views taken by the Hubble Space Telescope of a highly active region of star birth in 30 Doradus. The orientation and scale are identical for both views. Seven very young objects are identified with numbered arrows in the infrared image. Numbers 1 and 5 are new-born, compact clusters seen in visible light as well as the infrared image. Numbers 2 and 3 are new born stars or stellar systems still immersed within their *natal* dust and can be seen only as very faint, red points in the visible-light image. Number 4 is a very red star that has just formed within one of several very compact dust clouds nearby. Numbers 6 and 7 have been interpreted as 'impact points' produced by twin jets of material (see Section 5.3.4), originating from one of the young stars in number 5, slamming into surrounding dust clouds. The jets may be rotating anti-clockwise, thus producing moving, luminous patches on the surrounding dust, like a searchlight creating spots on clouds. (N. Walborn (STScI)/ R. Barba (La Plata Observatory)/NASA)

Another strand of evidence is that some dense clouds are found to contain a large number of compact infrared sources. A good example is shown in Figure 5.8, which shows visual and infrared images of part of the 30 Doradus nebula. This nebula contains some very young stars, which are unobservable in visible light but detectable in the infrared part of the spectrum. This may be due to the absorption of light at visual wavelengths by the dust surrounding a hot source or may be because the source is cool and emits only in the infrared.

QUESTION 5.2

Why should we expect dust clouds surrounding stars in the process of formation to emit in the infrared part of the spectrum? (*Hint*: think of Wien's displacement law.)

It seems that the infrared radiation comes not from the young star itself but from the cocoon of dust still surrounding it (called a **cocoon nebula**). Heated to a temperature of a few hundred kelvin by the recently formed star, the warm dust re-radiates in the infrared part of the spectrum.

The theoretical models of stellar evolution that you will look at in more detail shortly all point to regions of low temperature and high density as being the most likely sources of starbirth. The dense clouds seem to fit the bill quite well.

By way of a brief introduction to these models, recall that, according to Isaac Newton's 17th-century formulation of the theory of gravity, each piece of matter attracts every other piece by the force of gravity. The Earth attracts the Moon, and the Moon attracts the Earth. The Earth attracts an apple, making it fall down from a tree (and the apple also attracts the Earth). One apple also attracts another apple hanging on an adjacent branch, but in this case both of the masses are so small that a very sensitive device would be needed to measure the force. Just as the force between adjacent apples on a tree can be neglected for most practical purposes, so can that between molecules in the Earth's atmosphere, since these forces are much smaller than that between a molecule in the atmosphere and the Earth itself. However, when we are dealing with the gas and dust clouds spread through the vast expanses of the interstellar medium, we cannot ignore this force for two reasons. First, there are generally no other nearby large masses which dominate, and second, although the mass of each molecule or dust grain is very small, there are so many in an interstellar cloud that its total mass can be very large. Each molecule is affected by the gravitational attraction of the combined mass of all the others. James Jeans was able to show that, under appropriate conditions, a cloud (or part of one) would start to contract under the influence of the gravitational force. He derived a formula for calculating the mass and size that a cloud would have to reach, as a function of its temperature and density, before gravitational contraction could start. It is the details of this and subsequent processes, leading to the formation of stars, that are the subject of the remainder of this chapter.

5.3 Starbirth

5.3.1 Contraction of a dense cloud

All atoms, molecules and particles in a cloud are attracted to each other by gravitational forces. However, observations show that many clouds appear to be in a state of equilibrium – in other words, they don't seem to be contracting. Why is it then that the particles don't all collapse into a very small volume?

- Can you suggest a possible force to oppose gravitational contraction?

- Each gas particle (atom, molecule or free electron) is in continuous motion (with an average translational kinetic energy of $\frac{3}{2}kT$, where T is the temperature and k is the Boltzmann constant; Equation 4.2). This motion produces a gas pressure which provides an outward force to counteract the tendency of the gas to contract.

The basis of the approach used by Jeans was to consider the balance between the two forces. He proposed that if the force due to gravity was the greater, then gravitational contraction could occur. Using this simple criterion, Jeans was able to show that, for a given set of conditions of temperature and particle number density, there is a value for the mass of a uniform spherical cloud above which the force of gravitational attraction will overcome the opposing pressure due to the motion of the particles, and contraction will occur. This critical mass is known as the **Jeans mass** and is given by the following expression:

$$M_{\mathrm{J}} = \frac{9}{4} \times \left(\frac{1}{2\pi n}\right)^{1/2} \times \frac{1}{m^2} \times \left(\frac{kT}{G}\right)^{3/2} \tag{5.1}$$

where n is the particle number density, m the mass of the 'average' gas particle in the cloud, and T the gas temperature. Figure 5.9 shows how the Jeans mass, M_{J}, varies with temperature, T, and number density, n, when the particle mass, m, equals that of a hydrogen molecule. If the actual mass of a cloud exceeds M_{J}, then contraction is predicted.

QUESTION 5.3

According to the Jeans criterion, Equation 5.1, (or alternatively from Figure 5.9), what conditions of temperature, T, and number density, n, are likely to promote gravitational contraction? (A qualitative answer only is required.)

You have seen that the dense clouds are the densest and coolest regions in the interstellar medium, and so, by the Jeans criterion, would contract at lower mass than other types of region. Figure 5.9 shows that dense clouds of only a few solar masses would contract. Many dense clouds are far more massive than this. Moreover, there are yet denser regions within dense clouds (called cores and clumps), with masses between about $0.3M_{\odot}$ and about $10^3M_{\odot}$, that can therefore satisfy the Jeans criterion on their own. Thus, gravitational contraction of dense clouds is to be expected.

The picture that has been painted so far, however, is a highly simplified one.

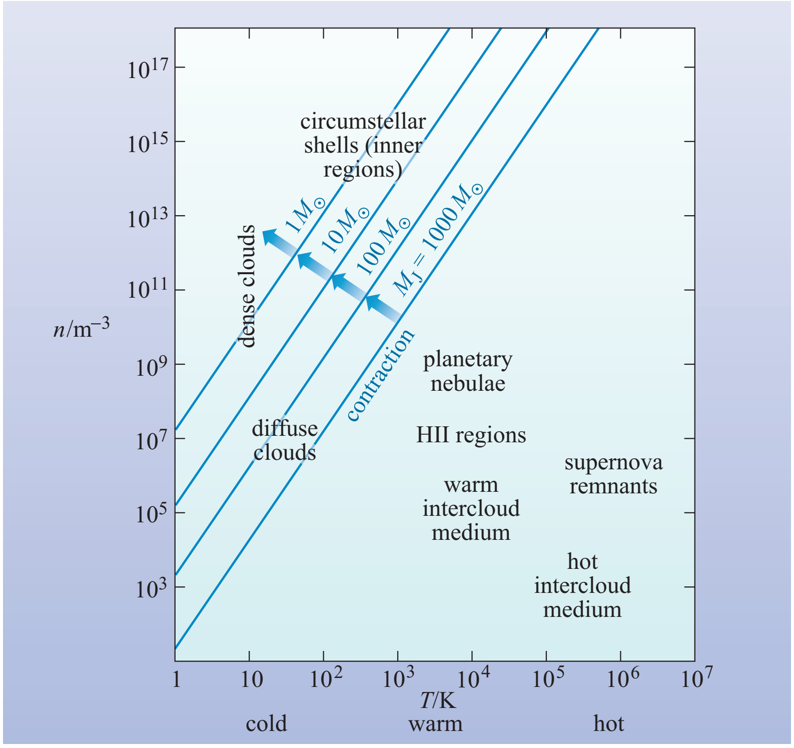

Figure 5.9 The relationship between temperature, T, particle number density, n, and the mass M_J, (the Jeans mass given by Equation 5.1) required in a spherical cloud for the gravitational force to cause contraction to occur. Any cloud or cloud fragment with a mass M_J which lies to the upper left of the diagonal lines will contract. It is assumed that the particles are hydrogen molecules, so $m = 2 \times 1.67 \times 10^{-27}$ kg.

■ What are the factors that could complicate this simplified approach to the gravitational contraction of a dense cloud?

❑ We wouldn't necessarily expect a typical cloud to be spherical or to have the same temperature and density throughout. In addition, no account has been taken of rotation or of magnetic fields, which almost certainly play an important part. These two factors may well inhibit a cloud's tendency to contract under gravity.

Any internal energy source will also act to inhibit contraction. In stars, which are the end result of contraction of a dense cloud, an energy source is provided by nuclear fusion. In Chapter 6 you will learn more about how stars are supported against gravitational collapse.

Despite its over-simplifications, the Jeans approach is a good starting point for more sophisticated, and perhaps more realistic treatments of the early stages of star formation. Whatever approach is adopted, we can speculate as to what might be the **trigger mechanism** that is probably needed to cause a cloud, or a region inside it, to change from an equilibrium state to one in which contraction has been initiated. Diffuse clouds do not satisfy the Jeans criterion for collapse but are often close to it, so any mechanism that pushes them over the limit acts like a trigger for collapse. This generally requires an increase in density (n in Equation 5.1) without a commensurately large increase in temperature. There are several mechanisms that have been suggested, and each appears plausible:

- Supernovae, about which you will learn more in Chapter 8, produce prodigious amounts of energy, some of which is carried away by a rapidly expanding shell of gas. This fast moving gas sweeps up the local interstellar material in front of it, producing a **shock front** or shock wave, a region of compressed gas immediately in front of the expanding shell. (This is a similar process to that

described for bow shocks, see Section 2.4.2 and 2.4.3.) Shock fronts alone are probably not sufficient to *produce* clouds but because they can gather matter *and* increase its density, and can also leave it in a cool state, they are particularly effective at making matter susceptible to Jeans contraction.

- Any shock will affect interstellar matter in this way: it does not have to be from a supernova. Another source of shock, though over a smaller volume of space than that from a supernova, is from regions where several O and B stars form – such groups of stars are called **OB associations**. Young O and B stars, which lie near the top of the main sequence, are massive, and highly luminous. The formation of several such stars close together produces large amounts of visible and ultraviolet light which causes a shock in the material in the surrounding ISM (due to the force of radiation pressure, see Section 5.3.4).
- The spiral structure of our Galaxy cannot be due simply to the different orbital speeds expected of material at different distances from the centre, as the spiral arms would 'wind up' in a relatively short time compared with its age. In order to maintain the spiral structure, a so-called **spiral density wave** sweeps around the Galaxy, compressing all the material that it passes, including clouds. In fact, the speed of these density waves is lower than the orbital speeds so the material is compressed as it sweeps through the density wave. Regions of star formation appear to be concentrated in spiral arms observed in other galaxies. This is consistent with the view that the density wave triggers Jeans collapse and forms stars.
- The close approach, or collision, of another cloud, or even a star, may be sufficient to produce a local gravitational disturbance that could trigger gravitational contraction.
- Star formation can also be triggered throughout a whole galaxy (called a starburst galaxy) by the interaction with another nearby galaxy.

Given that, somehow, a dense cloud, or a region inside it, starts to contract, we must now consider the question of what happens next.

5.3.2 Fragmentation

You have seen in Chapter 3 that stars are often found in groups, referred to as clusters. You also learnt that the stars in a cluster appear to have formed at about the same time. How can this be consistent with the picture of gravitational contraction that we've painted so far?

The answer is believed to lie in the phenomenon of **fragmentation**. As a dense cloud contracts then the density of the cloud (i.e. n) increases.

■ How will the Jeans mass change if n increases but all other quantities remain constant?

❑ From Equation 5.1, you can see that an increase in n will cause the Jeans mass, M_J, to decrease.

It is therefore possible that smaller parts of a massive cloud, which did not initially satisfy the Jeans criterion, can do so after the massive cloud has started to contract. Although the causes of fragmentation are unclear (it may be a result either of the initial 'clumpiness' of the cloud, or of rotation) it certainly appears consistent with the observation that star clusters are common. A cloud, initially with a mass of hundreds or even thousands of solar masses, can ultimately produce a large number

of small fragments, each collapsing on its own, to yield a cluster of stars. Star clusters formed in this way are called open clusters (see Section 3.2.4), reflecting their open structure (e.g. Figure 3.14). They typically contain a few hundred stars.

5.3.3 From a fragment to a protostar

We can better understand the evolution of a contracting cloud fragment by looking at the energy balance. Consider a single gas molecule in a contracting fragment. Initially, it will possess both gravitational potential energy and thermal energy (as translational kinetic energy of the molecule). For a molecule near the surface of the fragment, its gravitational energy is given by

$$E_g = -GMm/R \tag{5.2}$$

where M is the total mass of the fragment, m is the mass of the molecule, R is the radius of the fragment and G is the universal gravitational constant. Its translational kinetic energy is given by $E_k = \frac{3}{2}kT$ (Equation 4.2) where T is the temperature and k is the Boltzmann constant.

- ■ Assuming that the contracting fragment can be treated as an isolated system, what can we say about the total energy of the fragment as contraction progresses?
- ❑ The law of conservation of energy tells us that the total energy of an isolated system remains constant.

What does this mean for the contracting fragment? As contraction continues, the distance R of our molecule from the centre of the fragment will decrease. This results in a decrease in gravitational potential energy (if you are surprised by this, remember that there is a minus sign in Equation 5.2 for the gravitational energy). Because energy must be conserved, this reduction in gravitational potential energy is accompanied by an increase in other types of energy. The gravitational potential energy is converted into the molecule's kinetic energy. The molecules collide, and so the increase in the individual kinetic energies can be expressed as an increase in the thermal energy, $\frac{3}{2}kT$. This means that the temperature near the cloud surface will increase. Had we considered a molecule inside the cloud the expression for its gravitational energy would have been slightly different, but we would have reached the same conclusion.

Overall, therefore, as the fragment contracts, the gravitational energy of the particles is converted into the translational kinetic energy of molecules, which in turn is converted by mutual collisions into thermal energy of the gas and the temperature rises.

However, there are various complications. One is that collisions between molecules can leave them in excited states, which can emit characteristic radiation. In this case, the radiation is most likely to be in the radio wave, microwave or infrared part of the spectrum. Initially, this radiation tends to escape from the collapsing cloud and the resultant overall rise in temperature is minimal – perhaps only from 10 K to 20 K! However, as the contraction progresses, the number density of the molecules increases and this makes it more difficult for the emitted radiation to escape; it tends to be trapped by the surrounding layers. In other words, the gas becomes opaque to the radiation and now the internal temperature can rise more rapidly.

So although during the collapse process some molecules will be excited to emit radiation, this will only temporarily slow down the inexorable rise in temperature of the cloud fragment.

Section 4.2 introduced you to the Hertzsprung–Russell diagram and the role it plays in our understanding of stars.

- Where does a contracting fragment of a dense cloud fall on the H–R diagram at the *beginning* of its life?
- If the H–R diagram is plotted as in Figure 4.5, we would expect the fragment to be at the bottom right because of its low temperature and, before significant contraction has started, low luminosity.

The track of the contracting cloud fragment across the H–R diagram is far from certain and is difficult to observe directly for two reasons. One is that the process is taking place behind a shield of gas and dust which effectively screens the fragment from view. The second reason is the subject of Question 5.4 (below). For both reasons, we are forced to fall back on theoretical calculations and computer models of this phase. These seem to show that, after only a few thousand years of gravitational contraction, the surface has heated up to between 2000 and 3000 K. The fragment is still quite large at this stage and therefore the luminosity (which also depends on the surface area) can be quite high (10–100 times its eventual luminosity as a star on the main sequence). The exact track depends on the balance between the increasing surface temperature, which tends to increase the luminosity, and the decreasing surface area, which has the opposite effect.

- Can you recall an equation that reflects this balance?
- Equation 3.9, $L \approx 4\pi R^2 \sigma T^4$.

At this stage, the chain of events has started that will lead the fragment, almost inevitably, to become a normal main sequence star. For this reason, we are now justified in calling the fragment a **protostar**.

5.3.4 From a protostar to the main sequence

Figure 5.10 shows the predicted tracks for protostars of various masses as they evolve towards the main sequence region on the H–R diagram. They are called **Hayashi tracks** after the Japanese astrophysicist Chushiro Hayashi (1920–) who was pre-eminent in studying the evolution of pre main sequence stars in the 1960s.

Figure 5.10 also shows the timescale for this early stage of a star's evolution (from when the fragment that becomes the protostar breaks away from the parent cloud). It's clear from this that the more massive the protostar, the quicker it reaches the main sequence. For example, a protostar of $15M_\odot$ takes only about 10^5 years to reach the main sequence, less than 1% of the time that it would take a protostar of $1M_\odot$ to reach the same stage. Virtually all fragments that ultimately become main sequence stars take less than about 10^8 years to pass through the protostar phase – rather short on the astronomical timescale.

QUESTION 5.4

What are the implications of this short timescale on our ability to observe this phase of stellar evolution?

The details of the tracks in Figure 5.10 depend on complex changes in the internal structure and the way in which energy is transported through the protostars as they

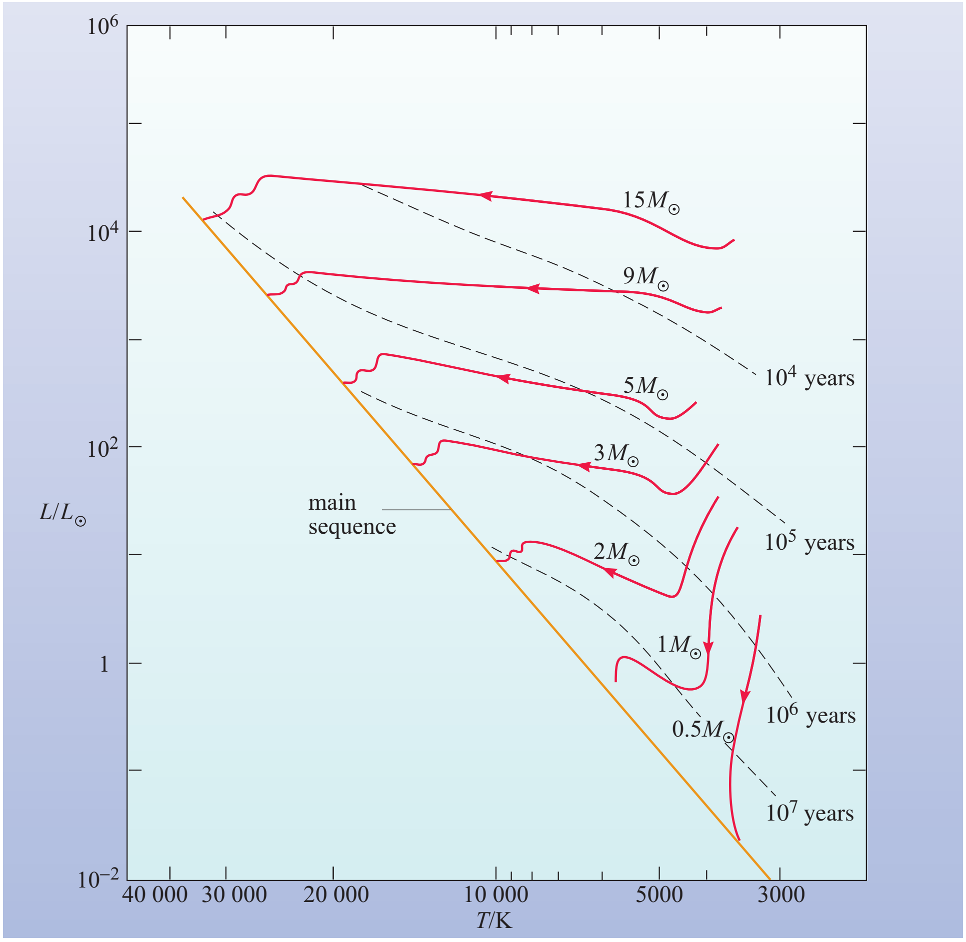

Figure 5.10 Theoretical tracks (Hayashi tracks) of protostars of various mass across the H–R diagram as they evolve towards the main sequence. Also shown are the times for the protostars to reach different stages of their evolution.

collapse. For protostars of intermediate and low mass, the early drop in luminosity is due to the effect of the increase in surface temperature T being more than offset by the effect of the decrease in radius R (remember Equation 3.9). Subsequently, for protostars more massive than about $2M_\odot$, the effects of increasing surface temperature and decreasing surface area just about balance so that the luminosity changes little as the temperature increases (indicated by an approximately horizontal track on the H–R diagram). Shortly before joining the main sequence, the tracks generally show a drop in luminosity as the effect of the contraction of the protostar tends to dominate over temperature effects.

Knowledge of the protostar phase of stellar evolution has been enhanced by observations made using radio telescopes. A large number of protostars show evidence of a phenomenon called **bipolar outflow** – that is gas flowing at high speeds, typically 50 km s^{-1}, in two streams moving in opposite directions. An example is shown in Figure 5.11, where the outflow shows up in the form of Doppler shifts in opposite directions, one a blue-shift and the other a red-shift (Section 3.2.1). Observations seem to indicate that the flows carry a significant amount of mass and require a lot of energy to sustain them – but they are believed to last for only a relatively short time, perhaps 10^4 years. As a cloud fragment contracts it spins faster and flattens into a **circumstellar disc** or torus with the protostar at the centre. If, at this stage, the protostar starts for some reason to produce a strong stellar wind – something that happens at various stages of a star's life, as we shall see later (Section 6.4.4) – then the disc will tend to channel the

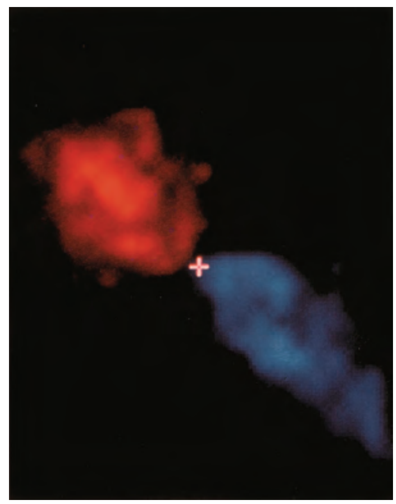

Figure 5.11 A millimetre-line image of the young stellar object IRS5, which lies in a dense cloud, catalogue number L1551. The image shows the intensity of the emission from CO molecules in the gas swept up by the material streaming away from the star (at the position marked by the cross). It has been colour-coded red where the material is moving away (i.e. a Doppler shift to longer wavelengths) and blue where the gas is approaching (i.e. a Doppler shift to shorter wavelengths). (R. Snell, University of Massachusetts)

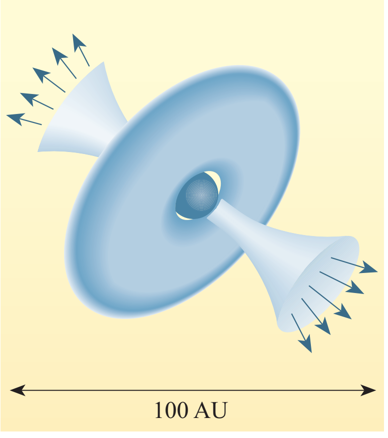

Figure 5.12 Schematic representation of bipolar outflow showing a central protostar, a circumstellar disc or torus and a strong stellar wind. The torus confines the wind to flow predominantly in two opposing directions.

outflowing material so that it streams out preferentially along the axis perpendicular to the disc in the form of two streams. The model as shown in Figure 5.12 may explain the principal observed features in bipolar sources, although current thinking suggests that the disc alone cannot channel the outflows, and that magnetic fields probably play a part too. Figure 5.13 illustrates examples of disc and jet structures, while in Figure 5.8 the effect of jets on the surrounding nebula can be seen in the infrared image (labelled 6 and 7, see caption for details).

One method for propelling stellar winds is the phenomenon called **radiation pressure**. It is a pressure exerted by photons on any object that absorbs or reflects them (see also Section 6.4.3). Although it is a weak force it is significant for individual atoms and molecules as well as small dust grains.

Figures 5.14 and 5.15 show nebulae within which individual protostars can be observed as the surrounding gas and dust clouds are eroded by another process – **photoevaporation**. Intense ultraviolet light from nearby highly luminous stars dissociates the hydrogen molecules (H_2) in the cloud into individual hydrogen atoms. The densest regions of the cloud survive longest and 'shield' material behind them forming the observed structures.

The winds, shocks, and UV radiation from young stars, particularly from O and B stars, are the main cause of disruption of dense clouds, and of any complexes of which they may be a part, and this disruption ends the process of star formation.

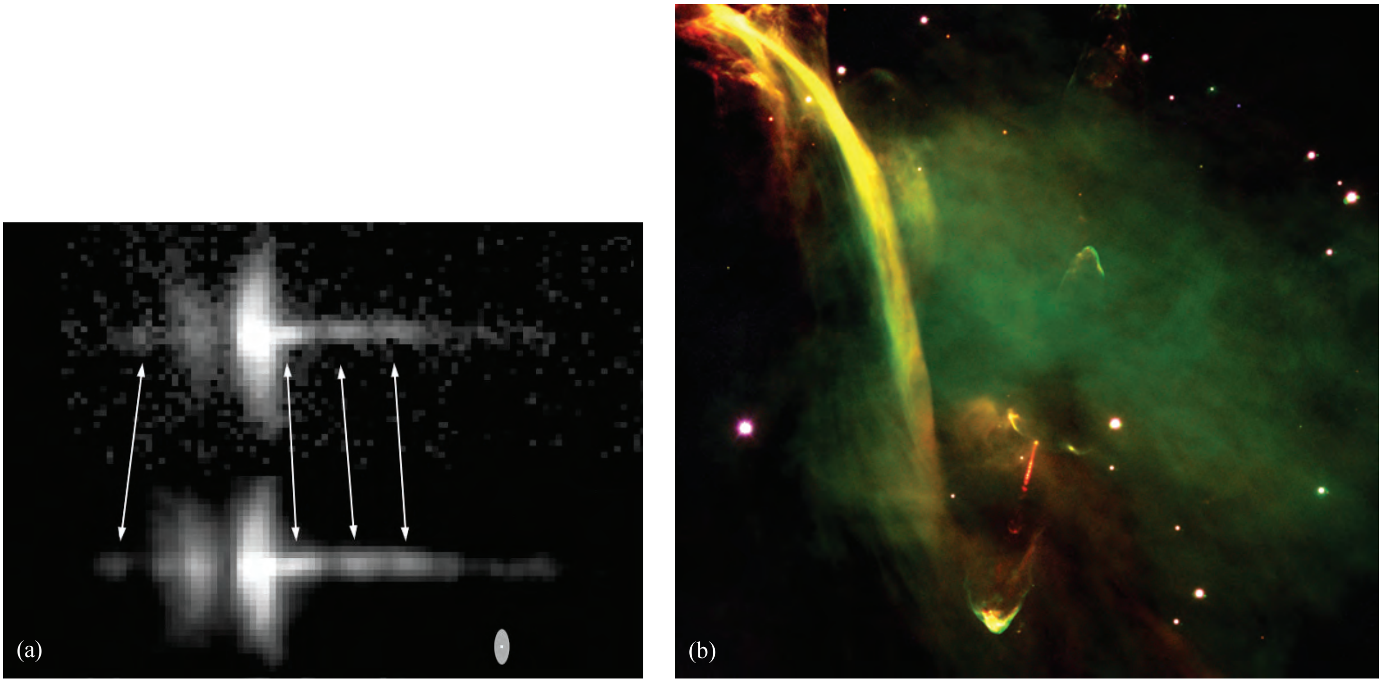

Figure 5.13 Images of protostellar objects illustrating the features shown in the schematic in Figure 5.12.
(a) Protostar HH30. The densest parts of the vertical edge-on disc obscure the newly forming star which illuminates the outside surfaces of the disc. The motion of blobs of gas within the jets indicated by the arrows can be seen in images taken a year apart. The small disc to the lower right indicates the size of Pluto's orbit around the Sun.
(b) Protostar HH34. The beaded structure of the jet (the red object in the lower centre of the image) indicates episodic outbursts of dense gas ejected when chunks of material fall onto the star from the surrounding disc. The effect of the two jets (the upper jet is obscured by intervening dust) is visible as they ram into the surrounding interstellar matter.
((a) C. Burrows (STScI)/NASA; (b) ESO)

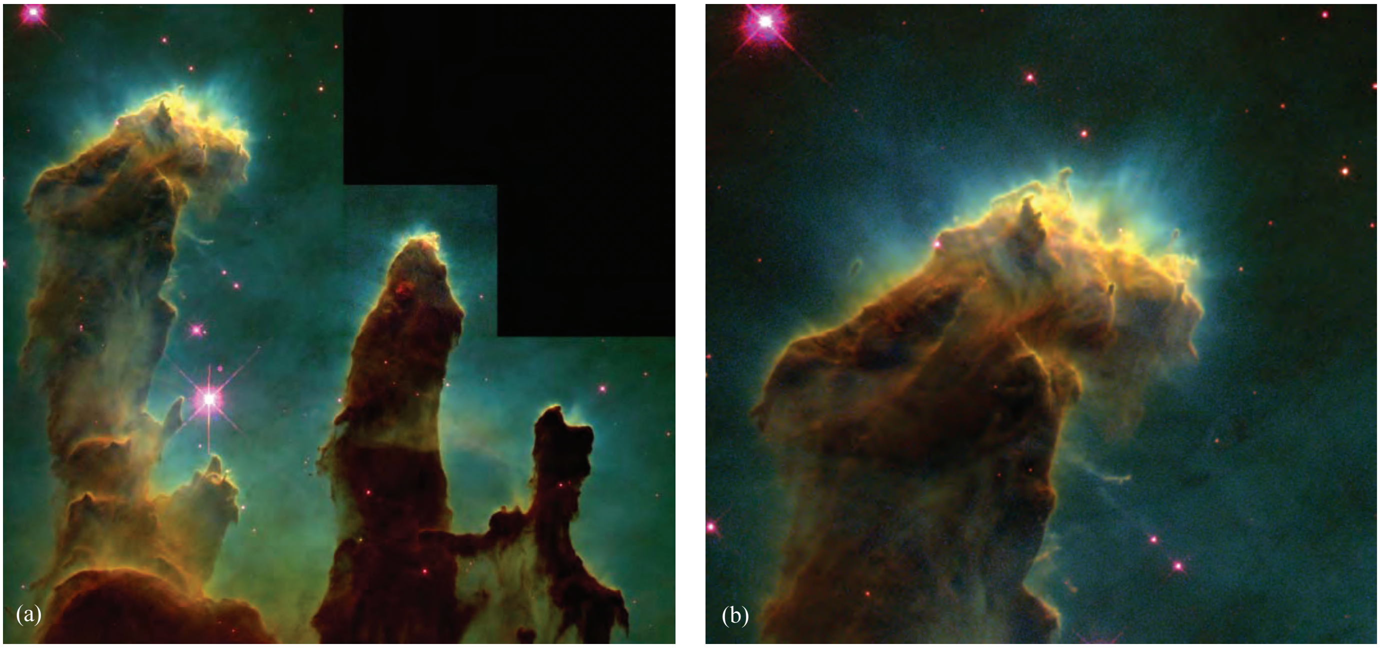

Figure 5.14 (a) One of the most famous images from the Hubble Space Telescope showing part of the Eagle Nebula in the constellation Serpens. The three pillars of dense gas and dust are being illuminated by newly formed massive stars that lie above the top of the image. The intense ultraviolet light is also eroding the nebula leaving columns behind the densest regions. The colour image is constructed from three separate images taken in the light of emission from different types of atoms; red from singly ionized sulfur atoms, green from hydrogen and blue from doubly ionized oxygen atoms. (b) This close-up shows 'fingers' of gas behind the cocoon nebulae containing protostars. In some cases the nebulae have themselves been eroded and the star inside revealed. (J. Hester and P. Scowen (Arizona State University)/NASA)

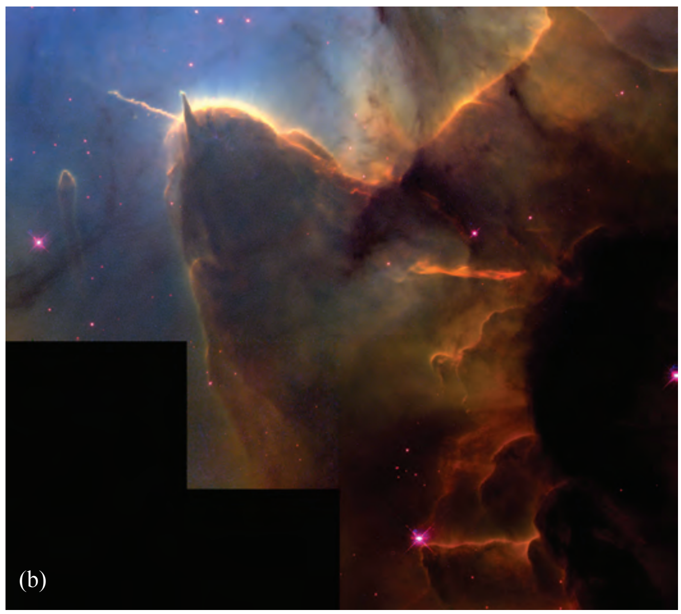

Figure 5.15 (a) The Trifid Nebula is a spectacular dense cloud in the constellation Sagittarius. The reddish colour is due to emission from hydrogen gas and the blue is due to starlight scattered by dust. (b) A Hubble Space Telescope view of a region 3 pc away from the nebula's central star, which is beyond the top of this picture. As in the case of the Eagle Nebula (Figure 5.14), this region is being eroded by the central star. A stellar jet (the thin, wispy object pointing to the upper left) protrudes from the head of a dense cloud and extends 0.2 pc into the nebula. The jet's source is a very young stellar object that lies buried within the cloud but which is likely to be exposed within the next 10 000 years. The nearby 'finger' pointing upwards shows a protostar which has already suffered this fate. ((a) NASA; (b) J. Hester (Arizona State University)/NASA)

■ Why would winds, shocks and UV radiation end the process of star formation?

❑ Extra energy sources in the collapsing cloud raise the temperature and hence raise the Jeans mass.

Only a small fraction of dense cloud mass is transformed into stars. The remnants of the dense cloud, and of any associated complex, can survive as independent clouds, perhaps to play a role in the build up of new, more massive clouds, and new giant molecular cloud complexes.

Another class of objects that is thought to be relevant to the early stages of stellar evolution is T Tauri stars, (which you have briefly met in Sections 3.3.5 and 4.2.3) named after the first to be discovered, the star designated 'T' in the constellation Taurus. They show several tell-tale signs of instability and youth, for example:

- They generally lie to the upper right of the main sequence on the H–R diagram, just where protostars are expected to lie;
- They usually appear in or near dense clouds;
- Many show an infrared excess (a higher flux in the infrared part of the spectrum than would be expected from a main sequence star at the appropriate temperature), suggesting a surrounding dust shell, probably in the process of being blown away;
- They often show irregular variability associated with strong magnetic fields that are not usually seen in older stars;
- They contain the element lithium that is normally destroyed later in the life of a star.

Other evidence indicates that they are young stars of age 10^5 to 10^8 years and that they are losing mass through stellar winds of high speed (up to 100–200 km s^{-1}). This is consistent with the previously discussed model of bipolar outflow sources that requires a strong stellar wind. It is thought that stars can lose as much as $0.5M_\odot$ in the form of a stellar wind during the T Tauri stage. Some T Tauri stars also show evidence of thin discs of circumstellar material, again consistent with the models of bipolar outflow sources, though in T Tauri stars the outflow is in all directions. From these observed properties, it is concluded that T Tauri stars are pre main sequence stars, in the mass range from about $0.2M_\odot$ to about $2.0M_\odot$, which are in their final spasmodic stage of birth and approaching the main sequence. If this interpretation is correct, they can be used as tracers of the behaviour of pre main sequence stars of these masses.

Let's attempt to bring together some of this observational evidence on the ever-useful H–R diagram. In Figure 5.16 a series of evolutionary tracks for pre main sequence stars of various masses is plotted. The line marked as 'birthline' represents the positions on the various Hayashi tracks where, according to one of the models, stars become optically visible. You may notice, incidentally, that these tracks differ in detail from those for a similar part of a star's life shown in Figure 5.10. This emphasizes that astronomers are not sure about the detailed evolution of a protostar onto the main sequence. Although the broad features are generally agreed on, the fine detail depends on various assumptions made in different models. The dots in Figure 5.16 mark the positions of observed T Tauri stars and stars that show evidence of major outflow (mostly in the form of bipolar outflow). They do indeed fall mostly in the region between the birthline and the main sequence. In addition, there seem to be more lying on Hayashi tracks for the lower masses, consistent with

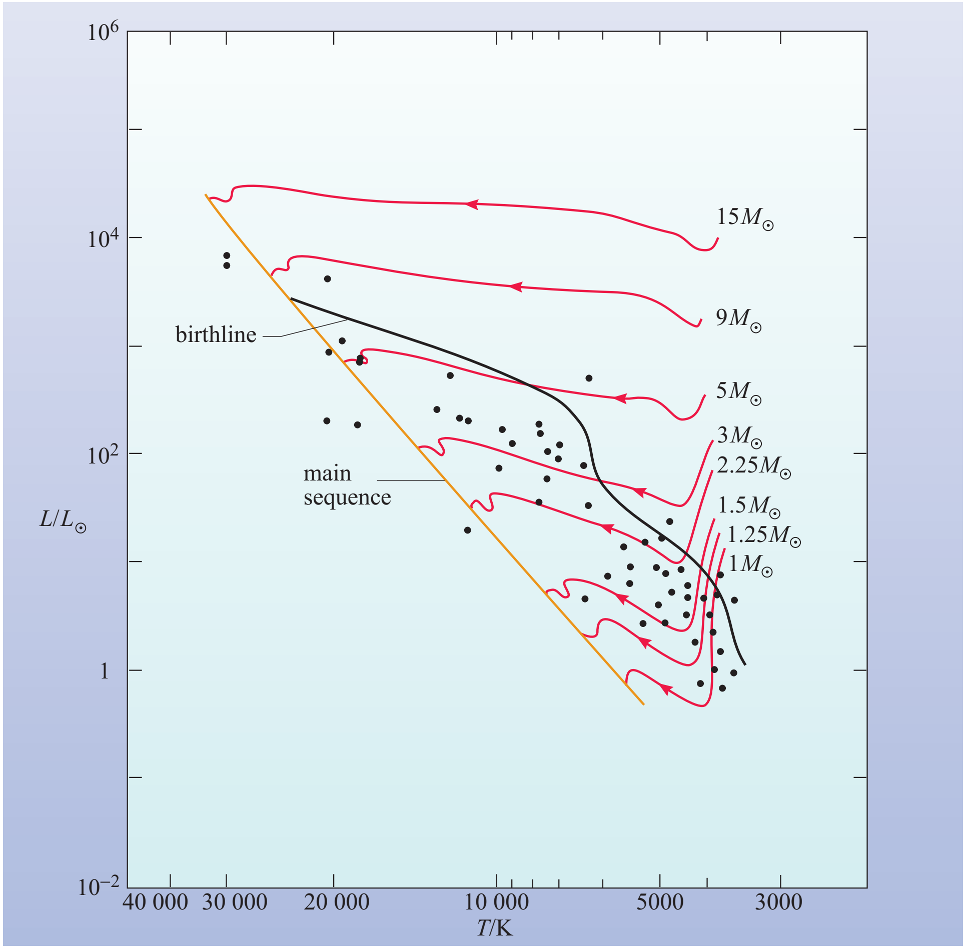

Figure 5.16 H–R diagram for a selection of T Tauri stars and stars showing evidence of major outflow.

our knowledge that low-mass stars are more common than high-mass stars. These observations confirm the general interpretation of both T Tauri stars and stars showing major outflow being associated with the early stages of stellar evolution.

QUESTION 5.5

What observational evidence from Doppler shifts can distinguish the mass outflow in a typical T Tauri star from that in bipolar outflow?

Although there are differences in the details of the various models of the later stages of collapse, they all essentially show a continued rise in temperature as the gravitational energy decreases. The critical point comes when the temperature in the centre, or core, of the protostar becomes sufficient for nuclear fusion to be triggered in the core. The energy released raises the core temperature sufficiently to halt contraction, and marks the protostar's arrival as a new main sequence star. You learned in Section 2.2.4 that it is nuclear fusion that is the power source in the Sun, so you shouldn't be surprised to hear that this is so for all main sequence stars. Before we consider the life of stars on the main sequence, we will briefly look at another formation process, (vital for our existence!) which has occurred for at least a few stars, the formation of planets.

5.3.5 Formation of planetary systems

The suggestion that discs of material may be associated with bipolar outflow sources and with some T Tauri stars posed the exciting possibility that this might also be material from which a planetary system may form, and that planetary formation might therefore be a widespread phenomenon in our Galaxy. Little is known about the formation of planets outside our own Solar System, their frequency, mass and orbital radius distributions. The current emphasis is on discovering other planetary systems and gathering evidence for their formation.

Observations of star-forming regions such as the Orion Nebula have revealed concentrations of dust around very young stars (see Figure 5.17). These discs are called protoplanetary discs, or proplyds, because it is believed that planets will eventually form from them. They are around ten times larger than our own Solar System and contain several Earth masses of dust.

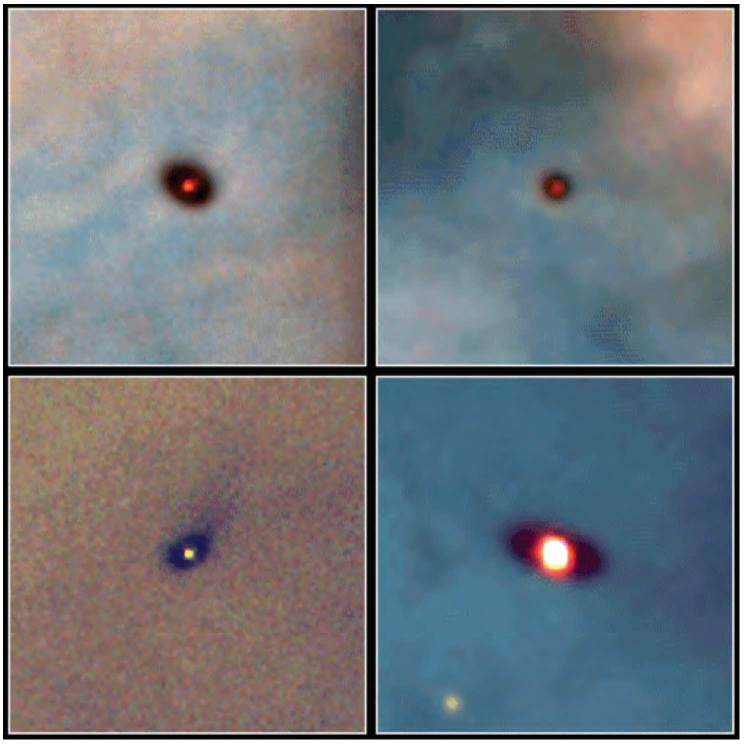

Figure 5.17 Protoplanetary discs in the Orion Nebula. The discs, surrounding the cool red central stars are seen in silhouette against the bright nebula, in these visible light images. (M. J. McCaughrean, (MPIA)/C. R. O'Dell (Vanderbilt University)/NASA)

In the last twenty years dust discs have been discovered around main sequence stars, and although it appears that these discs are not planetary systems in formation, they do provide some evidence of the presence of large objects in orbit.

Astronomers using the Infrared Astronomical Satellite (IRAS) in 1983 discovered that the spectra of various stars showed an infrared excess, probably because the stars are surrounded by a shell or disc of solid particles or dust. This material absorbs radiation from the central star and re-radiates at infrared wavelengths corresponding to the lower temperature ($T < 100$ K) of the dust. For some stars, the observations implied that the dust was in the form of a disc rather than a shell, giving support to the belief that what was being observed was the early stages of planetary system formation. The unexpected discovery by IRAS of an infrared excess from the bright star Vega ('vee-ga') was interpreted as emission from dust grains, about a thousand times larger than typical interstellar dust grains, in a shell or disc with a diameter of roughly 170 AU (about twice the diameter of our Solar System). Although Vega is a relatively

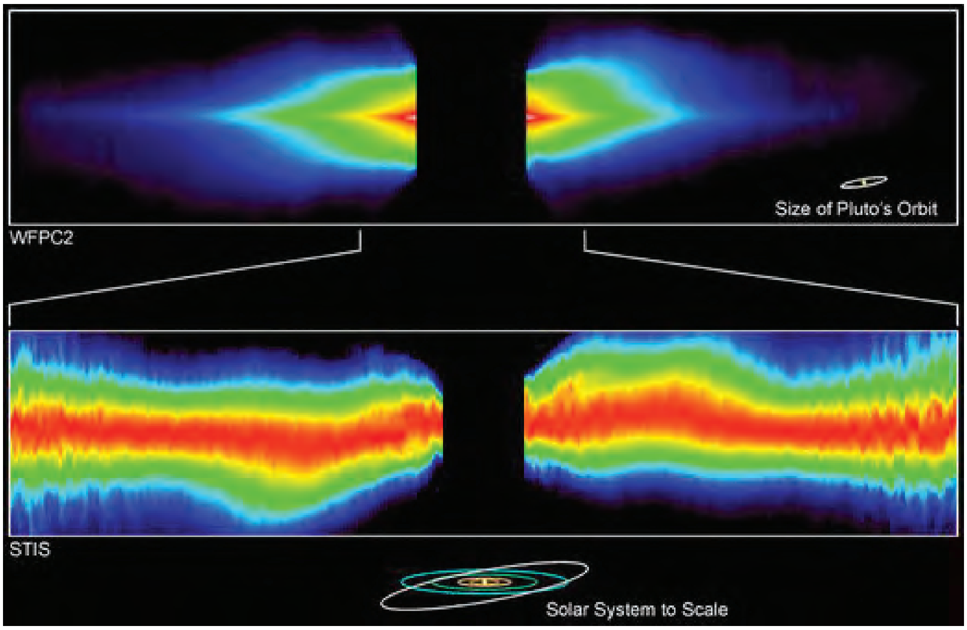

Figure 5.18 The upper image shows the central part of the disc of the young star β Pictoris, which is only a few hundred million years old, imaged at visible wavelengths using the Hubble Space Telescope. The black strip, much larger than the star, prevents the light from the star swamping the much feebler radiation from the disc. In the lower panel the intensity levels have been represented by false colours to enhance the warped shape of the disc believed to be due to the presence of one or more planets orbiting the star. (A. Schultz (Computer Sciences Corp.)/S. Heap (NASA GSFC)/NASA)

young star, perhaps only 20% of the Sun's age, it is probably still too old for dust to have survived from its formation. The disc may be a result of the break up of larger objects such as comets, but is still consistent with the suggestion that at least some of the process of planetary system formation has occurred.

Figure 5.18 shows an image of the region close to the star β Pictoris. The light from the star itself has been blocked out as it would otherwise swamp the much fainter light from the disc, which appears to extend out to several hundred AU from the star. The disc appears to be warped and has an inner edge (not seen in these images) implying the presence of one or more large unseen planets. The possible presence of planets, influencing the distribution of material in dust discs is even more apparent in the examples shown in Figure 5.19.

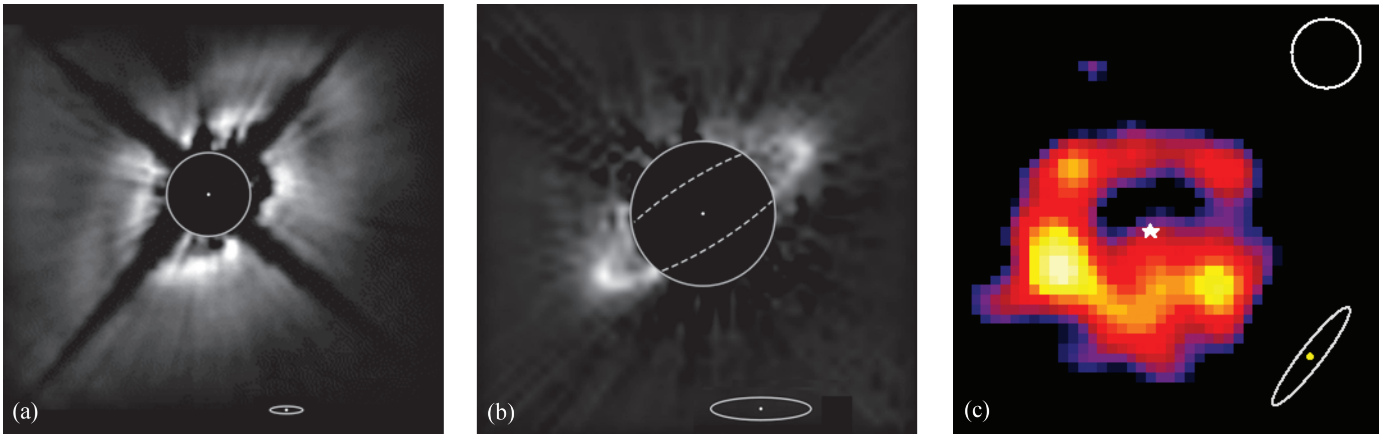

Figure 5.19 Images of circumstellar dust that suggest the presence of planets. The orbit of Pluto about the Sun is shown to the same scale. (a) The young star HD 141569 in the constellation Libra has a bright inner disc separated from a fainter outer disc by a gap which may have been carved out by an unseen planet.

(b) HD 4796A has a dust ring which can only stay intact by some mechanism confining the dust. The gravitational influence of unseen planets may be preventing the dust from dissipating due to collisions and radiation pressure. ((a), (b) B. Smith (University of Hawaii)/G. Schneider (University of Arizona)/NASA)

(c) Microwave image of a dust disc around the nearby star ε Eridani taken using the James Clark Maxwell Telescope. The irregular distribution of dust may indicate the presence of planets. The circle to the top right indicates the size of the telescope detector beam. ((c) Joint Astronomy Center, Hawaii)

Direct observation of planets is technically challenging due to their extreme faintness and close proximity to a much brighter star. However, it is possible to detect planets due to their effect on the motion of the star. In 1995 the first direct evidence of a body of planetary mass orbiting another main sequence star was obtained from observations of minute oscillations in the wavelength of spectral lines from the star 51 Pegasi with a period of 4.2 days. The Doppler shift of the lines was due to the orbital motion of the star about the common centre of mass of the star–planet system. This allowed a lower limit to the mass of the planet to be determined in a similar way to that for spectroscopic binary stars (see Section 3.3.7). The Doppler shift implied a planet of mass ≥ 0.47 times the mass of Jupiter (the exact mass is unknown because the tilt of the orbit to the line of sight cannot be determined by this method). Since then, many other stars have been found to have companions with masses comparable to, or a few times larger than, Jupiter. The majority of such objects have large masses but orbit very close to the parent star sometimes in highly eccentric orbits (unlike our Solar System where the massive gas giant planets formed in cool regions in the outer Solar System in near-circular orbits). It is possible that a similar formation process occurred as in our Solar System, with massive planets forming far from the stars, but subsequently undergoing interactions with remaining disc material or other protoplanets, with the result that they migrated inwards towards the central star. Another possibility is that some of these objects may not be planets at all but more massive brown dwarfs (see Section 6.4.2). It is certainly true that planets with large masses close to the primary star are easiest to detect.

- ■ Why are planets with larger masses close to a star easier to detect?
- ❑ Planets with larger masses and small distances exert a greater gravitational force on the star and therefore induce a greater 'wobble' of the star (corresponding to its orbit about the centre of mass of the star–planet system). This produces a greater Doppler shift in its spectral lines, which is more likely to be detectable, compared with a low-mass or distant planet.

The apparent scarcity of Jupiter-like planets in Jupiter-like orbits around other stars may therefore simply be a consequence of the fact that they are more difficult to detect combined with the short time since observations started.

The study of **extrasolar planets** is a rapidly advancing field in astronomy. One objective of such research is to determine whether the formation of planetary systems is commonplace. If our Solar System is typical, then there may be many planets capable of sustaining life elsewhere in the Galaxy. The next big step is to detect such Earth-like planets (and hence those which may harbour life) around other stars. Such low-mass planets will be difficult to detect using the spectroscopic method described above, but spectroscopy is likely to provide the solution in a different way. If a planet with an atmosphere passes in front of a star as it orbits, as viewed from the Earth, then it is possible to detect the gases present by the faint absorption lines they produce in the star's spectrum. It will therefore be possible to determine if the conditions for life, including the presence of oxygen, abundant in the Earth's atmosphere due to the respiration of plants, are present. Such techniques are already being attempted using ground-based telescopes. In the future, space observatories could provide a census of planetary systems in our neighbourhood of the Galaxy. The Gaia mission – see Section 3.2.2) will make precise measurements of the motions of the billion brightest objects in the sky, which is estimated to include 50 000 stars in our Galaxy possessing planets. Darwin, a possible future ESA mission, will make spectroscopic observations to search for Earth-like planets.

5.4 Summary of Chapter 5

- Stellar evolution occurs on such a long timescale that it can only rarely be directly observed.

The interstellar medium

- Stars form from material in the interstellar medium (ISM). The ISM contains regions with enormous variations of size, temperature and number density of particles.
- Even the densest clouds are far more tenuous than the Earth's atmosphere.
- Observational and theoretical evidence points to dense interstellar clouds as being the place where star formation begins.

Starbirth

- The Jeans criterion, despite being a simplified approach, makes useful predictions of the mass that a cloud, or part of one, must achieve before gravitational forces overcome those due to the thermal energy of the gas and contraction is able to start. The Jeans mass is given by

$$M_{\mathrm{J}} = \frac{9}{4} \times \left(\frac{1}{2\pi n}\right)^{1/2} \times \frac{1}{m^2} \times \left(\frac{kT}{G}\right)^{3/2} \tag{5.1}$$

  An external trigger mechanism is believed to cause a cloud to start contracting under the influence of gravitational forces. Compression of the cloud by shocks generated by supernovae and OB associations or by a density wave in spiral arms may be that trigger.
- As contraction of a dense cloud continues, the cloud fragments into smaller parts, each of which may continue to contract, as long as the Jeans criterion is satisfied for that particular fragment.
- Gravitational contraction is accompanied by a rise in the temperature throughout the fragment, though this is moderated by the escape of radiation from the fragment, particularly until the fragment becomes opaque.
- The temperature continues to increase and the size to decrease, both at a rate and in a way that depend predominantly on the mass of the fragment – now a protostar.
- Some protostars show evidence for circumstellar discs and for bipolar outflows, in which material flows out in opposite directions at high speed.
- T Tauri stars are pre main sequence stars of mass below about $2M_{\odot}$, showing strong stellar winds and variations in luminosity.
- When the temperature in the core of the protostar rises sufficiently, nuclear reactions are triggered. This provides the energy source to prevent further contraction, and at this stage the protostar joins the main sequence and becomes a fully fledged star. The time for a fragment to reach this stage is generally less than about 10^8 years; the more massive the fragment, the shorter the time.
- The presence of circumstellar discs around young stars led to the belief that they may be the precursers or remnants of planetary formation. The presence of planets around other stars has been inferred from the tiny *wobble* in the motion of those stars about the centre of mass of the star–planet system. The search for Earth-like planets and the conditions for life to form is currently one of the most rapidly expanding branches of astronomy.

Questions

QUESTION 5.6

A particular spherical dense cloud has a radius of 3 pc and a particle number density of $10^9\,m^{-3}$. Calculate whether this cloud contains enough material for a star or stars to form from it. (Assume that all the material in the cloud is in the form of H_2 molecules.)

QUESTION 5.7

A dense cloud is compressed, such that a denser core of 1 solar mass inside it increases in density and temperature, as in the table below. Assuming that the core consists mostly of H_2 molecules, and by considering the Jeans mass in each case, discuss whether the core will contract to form a star.

	Uncompressed	Compressed
number density/m^{-3}	5×10^9	8×10^{11}
temperature/K	10	25